干旱区地理
ARID LAND GEOGRAPHY

doi:10.12118/j.issn.1000-6060.2019.05.17

基于模糊聚类的新疆典型高寒草原土壤pH值空间制图①

朱 磊[1,2]， 盛建东[1,2]， 贾宏涛[1,2]
（1 新疆农业大学草业与环境科学学院，新疆 乌鲁木齐 830052；
2 新疆土壤与植物生态过程实验室，新疆 乌鲁木齐 830052）

摘 要：准确、高效地掌握草原土壤属性的空间分布能够为草地资源境管理提供基础信息和参考依据。相比于传统土壤调查方法，基于模糊逻辑的土壤—环境推理能够提高野外采样效率和预测制图精度，被广泛应用于数字土壤制图。但由于土壤自身的空间变异性及其与环境条件间的非线性，现有推理模型的稳定性较低，尚未在高寒草原区进行应用。选择新疆巴音布鲁克典型亚高山草原地区约4 km^2 区域为研究区，以高程、坡度、坡向、沿剖面曲率、沿等高线曲率、地形湿度指数6个地形因子为土壤环境因子，采用模糊 C 均值聚类（Fuzzy C-means Clustering，FCM）方法对环境因子聚类，得到9个环境因子组合，并在隶属度值高的环境因子组合中心共设置18个典型点。运用土壤—环境推理方法模拟研究区表层土壤pH值空间分布，其变化范围在7.170～8.186之间。选取35个独立样本进行精度检验（均匀采样点16个，横截面采样点9个，垂直带采样点10个），模拟结果与实测值基本吻合，且基于模糊聚类和土壤—环境推理方法的模拟精度高于普通克里格法和反距离权重法。通过基于模糊逻辑和土壤—环境推理的数字土壤制图方法在小尺度区域的运用验证，结果表明基于典型点的采样方案能够快速、有效地对区域土壤属性进行空间模拟，该方法对于类似小尺度的研究区同样有效。

关 键 词： 模糊聚类； 目的性采样； pH； 巴音布鲁克；

草地生态系统是陆地生态系统重要的可再生自然资源，为人类提供气候调节、土壤保持、涵养水源和畜产品等重要的服务功能[1]。但目前，我国90%的可利用草地都出现了不同程度的退化[2]。已有研究表明草地退化不仅是地表植被的退化，更与土壤退化具有密切联系，植被退化与草地土壤退化互为因果[3]。因此，准确、高效地掌握草地土壤的质量及空间分布能够为草地资源境管理提供基础信息和参考依据，对科学、持续地利用草地资源具有重要积极作用。

传统的土壤属性空间信息通常是借助空间插值方法获取，常用的方法有普通克里格插值法（Ordinary Kriging，OK）、反距离权重插值法（Inverse Distance Weighted，IDW）、普通最小二乘法（Ordinary Least Squares，OLS）等，还有学者利用多元线性回归分析集成所有成土因子对土壤养分进行空间分布预测，已取得较多研究成果[4-6]。近年来，随着遥感、GIS、人工智能和模糊推理等技术的发展，数字土壤制图领域取得了很大的进展[7-13]。其中，基于模糊逻辑的土壤—环境相似性推理研究日益受到关注，最具代表性的是朱阿兴等提出的SoLIM模型（Soil-Land Inference Model），其利用人工智能获取并量化土壤和环境关系的知识，通过相似度推理，使土壤类型的推测精度有较大提高[14-16]。此外，部分学者运用相似性分布和土壤类型的典型值进行土壤属性定量模拟，所得土壤类型图精度均优于传统方法[13,17-24]。还有学者将土壤—环境推理与目的性采样相结合，提高了精

① **收稿日期**：2019-03-10； **修订日期**：2019-06-11
基金项目：国家自然科学基金项目（31560171）；新疆农业大学博士后基金联合资助
作者简介：朱磊（1982-），男，博士，副教授，研究方向地理信息系统与遥感应用、土壤属性空间模拟. E-mail：smartzhulei04@163.com

细土壤普查的效率和数字土壤制图的精度[19,20,22,25-26]。但由于土壤自身的空间变异性及其与环境条件间的高度非线性，现有推理模型的稳定性较低，难以在不同区域进行直接推广。

新疆巴音布鲁克草原位于天山南坡中部，是我国最大的荒漠区亚高山高寒草原，重要的畜牧生产基地[27-29]。区域内典型的高寒草原土壤发育特征及其环境要素具有鲜明的地域特点，而基于土壤—环境的协变关系开展草原土壤属性空间制图的定量研究还相对较少。因此，本研究在巴音布鲁克典型高寒草原内选择研究区，运用模糊聚类、GIS 和目的性采样方法，构建土壤—环境推理模型，以土壤 pH 值为例模拟区域土壤属性的空间分布，并与 OK 和 IDW 等传统方法进行比较。研究将有助于深入理解新疆高寒草原土壤发生、演变过程与环境要素的关系，丰富草地土壤学科的案例研究，同时可提高山区土壤调查的效率和精度。

1 研究区概况

研究区位于新疆巴音郭楞蒙古自治州和静县巴音布鲁克镇境内，面积约为 4 km^2，地势北高南低，具有相对平缓的山脊，较陡的山坡和宽平的沟谷，海拔在 2 603 ~ 2 995 m 之间（图 1）。该区域年平均气温 -4.8 ℃，1 月最低气温可达 -48 ℃，7 月最高气温 30.5 ℃，年平均风速 2.7 $m \cdot s^{-1}$，年降水量平均 276.2 mm，年蒸发量高达 1 022.9 ~ 1 247.5 mm，年日照 2 466 ~ 2 616 h，热能 562.8 $kJ \cdot cm^{-2} \cdot a^{-1}$，全年积雪日达 150 ~ 180 d，无绝对无霜期，属典型的高寒气候[30]。研究区土壤主要为亚高山草原土，地表植被覆盖度较高，以高寒草原类为主，主要包括紫花针茅（*Stipa purpurea*）、线叶蒿草（*Kobresia capillifolia*）和珠芽蓼（*Polygonum viviparum*）等。

2 研究方法

研究将首先辨识并空间量化区域内土壤—环境协变因子，建立基础要素数据集，采用模糊 *C* 均值聚类（Fuzzy C-means Clustering，FCM）方法提取环境组合，在各组合典型点野外采样获取表层土壤 pH 值，最终运用土壤—环境相似性推理方法实现研究区表层土壤 pH 空间分布制图，并与传统空间插值方法进行比较。

2.1 环境要素集的建立

在土壤发育过程中，气候、母质、时间、地形等因素都可以影响土壤的 pH 值。然而在景观和小流域尺度上，地形起到主要作用[10,31-34]，其通过影响母质和水热条件再分配[35]，进而导致 pH 值的空间变化。由于本研究区范围较小，气候和母质要素可视为均一，研究中未予考虑。此外，区域地质年龄的参数化难度较大，而半定量、基于样点的分析又难以形成具有足够精度的栅格数据[15]，故研究中尚不考虑

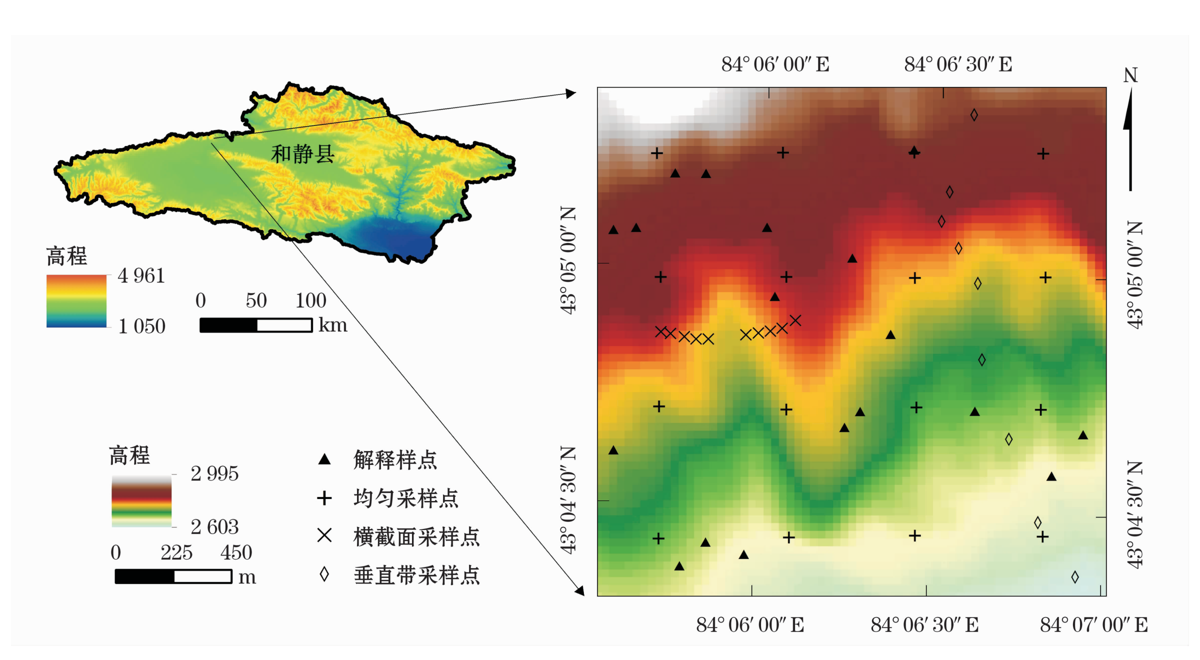


图 1 研究区位置及样点分布

Fig. 1 Location of the study area and field sampling points

时间因素。因此，本研究将区域地形特征作为土壤属性的主要协变因子，参照土壤专家的意见，选取高程、坡度、坡向、沿剖面曲率、沿等高线曲率和地形湿度指数 6 个因子构建土壤—环境因子空间数据集（图 2）。基础数据为分辨率为 30 m 的 DEM，其它因子在其基础上基于 ArcGIS10.0 平台计算而成。

2.2　基于模糊聚类的环境组合获取

采用 FCM 方法得到的环境因子组合代表了土壤属性在空间变化的不同模式，并以模糊隶属度的形式表达，根据隶属度值的高低判别环境因子组合中心位置以及组合间的过渡区[36]。该方法通过计算多属性空间中的距离，使模糊聚类达到最优，并以此确定各数据点的隶属度[15,37]。模糊目标函数如下：

$$J_m(U,v) = \sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^m d_{ik}^2 \qquad (1)$$

$$d_{ik}^2 = \| y_k - v_i \|_A^2 = (y_k - v_i)^t A (y_k - v_i) \quad y \in Y \qquad (2)$$

式中：U 为模糊聚类的隶属度矩阵；v 为聚类中心集；Y 为环境要素数据集；n 为 Y 中数据的个数；c 为聚类类别的个数；m 为加权指数；d_{ik}为数据点 y_k 到中心 v_i 的加权距离；u_{ik}为第 k 个数据点属于第 i 类的隶属度；A 为距离权重矩阵；J_m 为模糊分类误差。

FCM 通过分割系数（Partition coefficient，F）和标准化分类熵（Normalized entropy，H）两个指标随聚类数的变化情况来确定最优聚类数目[10,37]。随着聚类数的增加，当分割系数从（$c-1$）到 c 的递减量比从 c 到（$c+1$）的递减量小时，可认为此时的聚类结果较稳定，对应的聚类数为最佳[15]。

2.3　典型点土壤信息提取与土壤属性制图

将各环境因子组合中隶属度最大值所在位置作为该组合类的典型点，构成土壤与环境的解释集。野外 GPS 定位后，在各典型点采集土壤表层样本（0～20 cm），经风干、去杂、磨细、过筛等处理后，采用电位法测定土壤样本 pH 值，测定时土水比为 1:2.5，每个土样重复测定 3 次，结果取平均值[38－39]。

环境因子组合类代表了土壤属性在空间变化的不同模式，其模糊隶属度图可表达区域内各像素对其典型模式的相似程度。基于环境因子组合的模糊隶属度和典型点的土壤属性值，选择加权平均模型推理研究区表层土壤 pH 值。计算公式如下[40]：

$$V_{ij} = \sum_{k=1}^{n} S_{ij}^k V^k \Big/ \sum_{k=1}^{n} S_{ij}^k \qquad (3)$$

式中：V_{ij}为区域内未知像素点（i，j）的土壤属性值；V^k 为典型点 k 对应模式的典型土壤属性值；S_{ij}^k为像素（i，j）对应环境组合类的模糊隶属度值；n 为环境组合类总数。

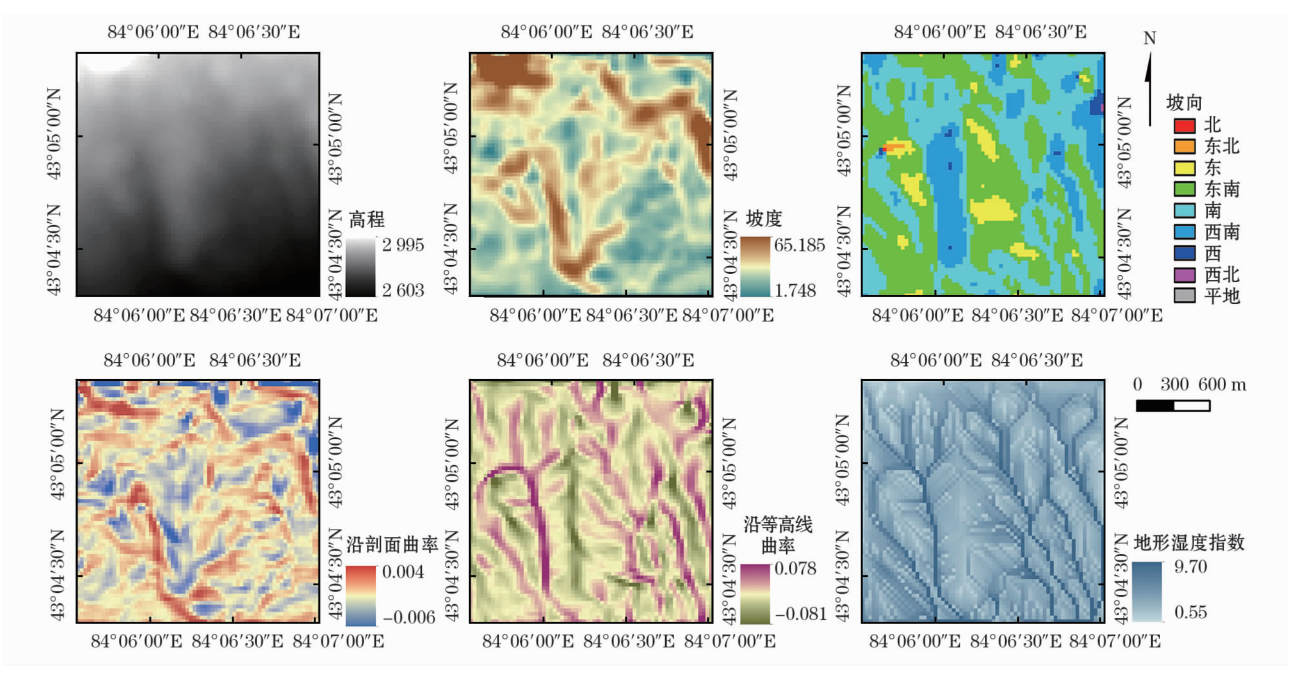


图 2　研究区环境因子分布

Fig. 2　Distribution of derived environment parameters of study area

2.4 精度检验

为了检验上述方法所得土壤 pH 空间分布图的精度,研究共采集 35 个独立样点建立了验证集(图 1)。其中,均匀采样点 16 个,按照 500 m×500 m 等间隔布置样点,用来检验区域土壤属性的整体模拟精度;横截面采样点 10 个,设计为横穿山坡、沟谷的线路,即通过采样点在较短的距离内穿越研究区主要的地貌类型变化,检验预测土壤 pH 图是否可以较好地反映土壤信息的微地形渐变特征;垂直带采样 9 个,以检验土壤 pH 值沿海拔梯度变化的模拟效果。验证样本的野外采集方法和实验室处理分析流程与解释集相同。选取平均绝对误差(Mean Absolute Error,*MAE*)、均方根误差(Root Mean Squared Error,*RMSE*)和一致性系数(Agreement Coefficient,*AC*)对推理结果进行定量评价,并与 OK 和 IDW 方法进行比较。

3 结 果

3.1 训练样本集构建

研究采用 5 种不同的模糊权重对数据层进行迭代计算(m=1.5,1.75,2,2.25,2.5)[14],各权重下均产生聚类数从 2 到 20 的不同聚类结果,基于公式(1)和(2)计算分割系数和熵随 n 增大产生的差值。结果显示,当 m=2、c=9 时 H 增幅最小且 F 变化较前后两类变化小且趋于稳定,故本研究中选择最佳类别数为 9,加权指数为 2,进而得到各环境因子聚类组合模糊隶属度分布。图 3 列举了 9 种组合中的 3 组,class1 代表坡脚和沟谷区域,class2 代表海拔相对较低、坡度较缓的平原部位,class3 代表海拔相对较高的坡肩部位。随后将各环境因子组合中隶属度值最大处的样点作为土壤属性推理的解释样点,构成解释集(图 1)。

3.2 表层土壤 pH 值推理制图

对典型样点集进行一般性统计,pH 值范围为 7.287 ~ 8.263,平均值为 7.595,变异系数为 3.32%,表明研究区土壤为典型的弱碱性土壤。根据土壤—环境的空间隶属度分布,结合环境要素数据库和实验室数据,由公式(3)得到研究区 0 ~ 20 cm 表层土壤 pH 值空间分布(图 4a)。结果显示研究区表层土壤整体呈现弱碱性,pH 值随地形连续变化,变化范围为 7.170 ~ 8.186。与采用 OK 和 IDW 的制图结果(图 4b、图 4c)相比,FCM 方法所预测的土壤 pH 分布图较好地反映出表层土壤 pH 对地形变化的响应,显示出研究区表层土壤 pH 空间分布的更多细节,对于土壤模拟结果更加接近实际。

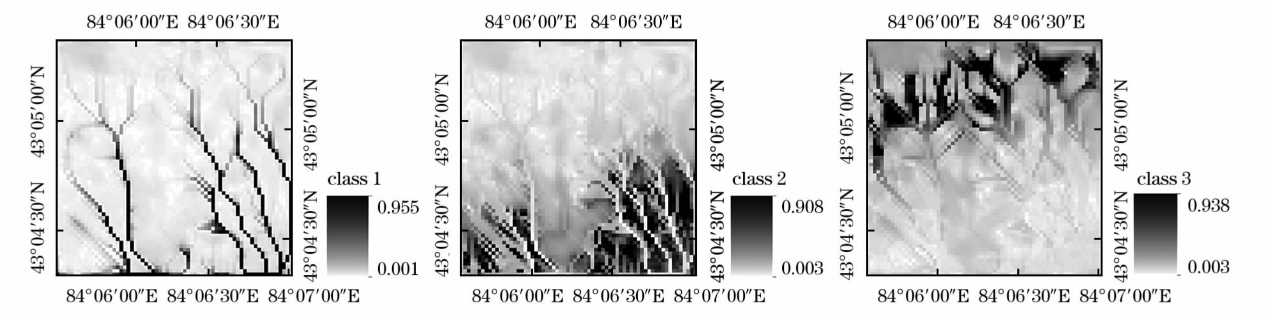


图 3 环境组合的模糊隶属度分布(以 class1、class2 和 class3 为例)

Fig. 3 Fuzzy membership maps of combinations of environment factors based on FCM(class1, class2 and class3)

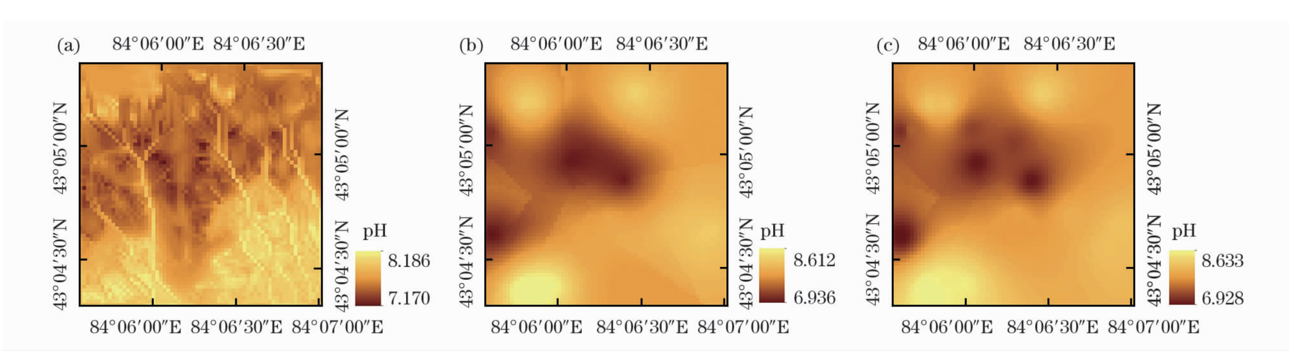


图 4 表层土壤 pH 值空间分布模拟结果(a-FCM,b-OK,c-IDW)

Fig. 4 Simulation results of spatial distribution of pH in surface soil(a-FCM,b-OK,c-IDW)

3.3　精度评价

结果显示,除个别样点偏差相对较大外,推理结果与实测结果总体基本吻合,距离聚类中心较近且地形变化较显著的区域,模拟精度相对较高。全部验证点的整体平均绝对误差和均方根误差分别为 0.14 和 0.19,不同采样方式表现出的模拟精度特征有所差异。横穿山坡、沟谷的线路的横截面采样点的模拟值与实测值最为接近;垂直带采样方式次之,均匀采样点的模拟精度相对最低。横截面采样点在较短的距离内穿越了研究区内地貌类型变化相对较大的山坡、沟谷,因此所建立的推测模型能更好的捕捉地形变化较大处的 pH 值变化,所以模拟精度相对较高。这也反映模型应用在环境梯度变化大的地方具有更好的有效性。

对比土壤—环境推理方法和其它两种方法的模拟精度,可以得出:OK 法和 IDW 法的 *MAE* 和 *RMSE* 较为接近,均高于 FCM 法,说明本研究所采用的土壤—环境推理方法所得预测结果精度高于其它两种方法(图 5 和表 1)。此外,将 3 种方法在验证点处的预测值与实测值进行比较,结果显示传统插值方法预测值的回归决定系数分别为 0.22 和0.29,低于 FCM 法的 0.47,可见 FCM 法的预测精度高于 OK 插值和 IDW 插值方法。同时,一致性系数显示土壤—环境推理方法的稳定性较高。

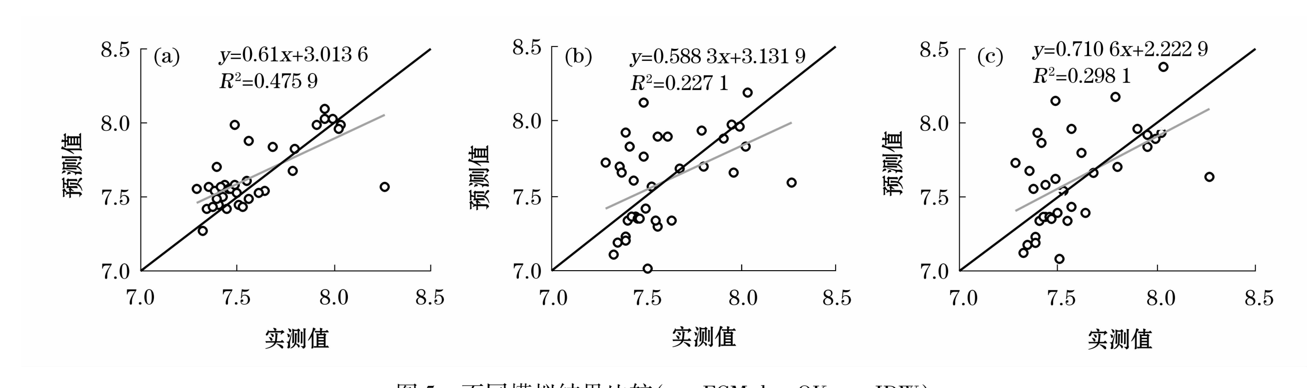


图 5　不同模拟结果比较(a－FCM,b－OK,c－IDW)

Fig. 5　Comparison of different simulation results(a－FCM,b－OK,c－IDW)

表 1　不同方法的精度评价结果对比

Tab. 1　Accuracy evaluation of different methods

模拟方法	验证点类型	*MAE*	*RMSE*	*AC*
FCM	总体	0.14	0.19	0.81
	均匀采样	0.33	0.26	
	横截面采样	0.08	0.09	
	垂直带采样	0.10	0.13	
OK	总体	0.23	0.29	0.68
	均匀采样	0.51	0.36	
	横截面采样	0.15	0.16	
	垂直带采样	0.22	0.25	
IDW	总体	0.22	0.28	0.72
	均匀采样	0.51	0.35	
	横截面采样	0.14	0.15	
	垂直带采样	0.18	0.23	

4　讨　论

目的性采样过程中,样点的选定是土壤属性模拟精度的重要前提。景观尺度下,决定土壤变化的主要因子是地形和水文状况[10],尤其在较小空间范围里的精细土壤调查中,地形要素可代表土壤形成与发展的主要影响因素[34]。由于研究区面积较小,且受限于数据的空间量化可能性与可获取性,本研究中对于少量野外样本的确定主要基于数字高程模型及其衍生出的地形因子,未将其它土壤环境因素考虑在内。如研究区扩大至流域和区域尺度时,气候因子、母质因子和人类活动因子则是模拟时必须要考虑的因素,将有助于样点典型性模型模拟精度的提高。

OK 法和 IDW 法是最为常用的两种空间插值方法,但它们只注重样点的空间位置和数据本身,未考虑环境因素的作用。OLS 法虽注重环境因素的作用,忽视了空间位置的影响[41]。近年来被广泛使用的地理权重回归法(Geographically Weighted Regression,GWR),对 OLS 方法进行了改进,同时考虑了插值点的空间位置和环境要素的作用,常被用于对土壤属性模拟和空间非稳定关系的探讨[42-44]。但其本质上是局部地理回归的结合,随着被估计的局部模型越来越多,显著性水平测试有意义的概率会越来越大。此外,不恰当的带宽参数会引起回归系数出现强烈的空间变异或平滑,被估计的回归系数的空间模式被人为地引入,影响实际回归效

果[43,45]。上述方法中样本的数量对模拟精度影响较大,当样本较少时,往往难以建立可靠的模型。基于土壤—环境协变关系和模糊聚类的建模方法针对典型环境类型进行目的性采样,能够在有限的样本数量下保证一定的模拟精度,因此在兼顾效率与精度时,具有相对的优越性。

基于 DEM 的数字地形分析,作为定量提取各种地形信息的有效手段,已被广泛应用于水文学、地貌学、土壤学、生态学等诸多研究领域[46]。DEM 分辨率的变化会直接影响地形分析的结果[47-48],在实际应用中可通过对比选择适宜的 DEM 分辨率[49]。本研究所用 DEM 数据的空间分辨率为 30 m,提取的各地形因子均在此基础上完成。但如采用不同分辨率的地形数据,在运用模糊聚类提取环境组合时会存在怎样的差异目前尚缺乏研究,今后可通过模拟精度比较分析得出土壤属性推理的适宜分辨率和最佳推理空间。此外,本研究属于小尺度的土壤属性模拟,现有研究中已有不少学者在不同空间尺度开展了相关研究[18-23],但土壤—环境推理的过程和结果是否具有尺度效应,如何进行多尺度推演等,有待深入研究。

5 结 论

本研究以新疆巴音布鲁克山区典型高寒草原为对象,选取高程、坡度、坡向、沿剖面曲率、沿等高线曲率、地形湿度指数等地形因子进行模糊聚类,获得研究区影响土壤属性的典型环境因子组合,结合目的性采样,运用土壤—环境关系模拟研究区表层土壤 pH 值空间分布,并与传统空间插值方法进行比较。结论如下:

(1) 基于构建的基础数据集,采用模糊 c 均值聚类方法得到了 9 类环境因子组合,并在各组合隶属度值最高的中心设置了 18 个典型点,运用土壤—环境关系模型模拟得到研究区表层土壤 pH 值空间分布图,区域变化范围 7.170 ~ 8.186 之间,整体呈弱碱性。

(2) 模糊聚类方法通过较少的典型样点获取了区域内土壤属性的连续性空间分布,提高了采样的有效性,模拟结果较好地反映了表层土壤 pH 对地形变化的响应,显示出研究区表层土壤 pH 空间分布的更多细节。

(3) 经 35 个独立样本检验,研究区土壤 pH 值模拟结果与实测值基本吻合,且模拟精度高于普通克里格法(OK)和反距离权重法(IDW),是一种相对低成本、高效率的数字土壤制图方法,在草地土壤调查中具有较好的应用潜力。

参考文献(References)

[1] 岳平,宋韦,李凯辉,等. 天山中部巴音布鲁克高寒草原大气无机氮沉降[J]. 应用生态学报,2014,25(6):1592-1598. [YUE Ping, SONG Wei, LI Kaihui, et al. Inorganic N deposition in the Bayinbuluk alpine grassland of the central Tianshan Mountains [J]. Chinese Journal of Applied Ecology, 2014, 25(6):1592-1598.]

[2] 张英俊. 草地退化原因及防治对策[C]//2007 中国草业发展论坛论文集, 2007:163-164. [ZHANG Yingjun. Causes and countermeasures of grassland degradation[C]//2007 China Grass Industry Development Forum, 2007:163-164.]

[3] 李建宏,李雪萍,卢虎,等. 高寒地区不同退化草地植被特性和土壤固氮菌群特性及其相关性[J]. 生态学报,2017,37(11):3647-3654. [LI Jianhong, LI Xueping, LU Hu, et al. Characteristics of and the correlation between, vegetation and N-fixing soil bacteria in alpine grassland showing various degrees of degradation [J]. Acta Ecologica Sinica, 2017, 37(11):3647-3654.]

[4] 李启权,岳天祥,范泽孟,等. 中国表层土壤全氮的空间模拟分析[J]. 地理研究,2010,29(11):1981-1992. [LI Qiquan, YUE Tianxiang, FAN Zemeng, et al. Spatial simulation of topsoil TN at the national scale in China [J]. Geographical Research, 2010, 29(11):1981-1992.]

[5] 刘杰,张杨珠,罗尊长,等. 湘南丘岗红壤地区土壤肥力质量指标的空间变异特征[J]. 湖南农业大学学报(自然科学版), 2012,38(1):96-101. [LIU Jie, ZHANG Yangzhu, LUO Zunchang, et al. Spatial variability of soil fertility quality indices in red soil region [J]. Journal of Hunan Agricultural University (Natural Science Edition), 2012, 38(1):96-101.]

[6] 黄安,杨联安,杜挺,等. 基于多元成土因素的土壤有机质空间分布分析[J]. 干旱区地理,2015,38(5):994-1003. [HUANG An, YANG Lian'an, DU Ting, et al. Spatial distribution of the soil organic matter based on multiple soil factors[J]. Arid Land Geography, 2015, 38(5):994-1003.]

[7] ZHU A X, HUDSON B, BURT J E, et al. Soil mapping using GIS, expert knowledge, and fuzzy logic[J]. Soil Science Society of American Journal, 2001, 65:1463-1472.

[8] MCBRATNEY A B, MENDONCA SANTOS M L, MINASNY B. On digital soil mapping[J]. Geoderma, 2003, 117:3-52.

[9] 张华,张甘霖,龚子同. 土壤-景观定量模型研究进展[J]. 土壤通报,2004,35(3):339-346. [ZHANG Hua, ZHANG Ganlin, GONG Zitong. The progress of quantitative soil-landscape modeling: A review[J]. Chinese Journal of Soil Science, 2004, 35(3):339-346.]

[10] 杨琳,朱阿兴,李宝林,等. 应用模糊 c 均值聚类获取土壤制图所需土壤-环境关系知识的方法研究[J]. 土壤学报,2007,44

(5):784 - 791. [YANG Lin,ZHU Axing,LI Baolin,et al. Extraction of knowledge about soil-environment relationship for soil mapping using fuzzy c-means(FCM)clustering[J]. Acta Pedologica Sinica,2007,44(5):784 - 791.]

[11] MINASNY B,MCBRATNEY A B,LARK R M. Chapter 2. Digital soil mapping technologies for countries with sparse data infrastructures[M]//HARTEMINK A E,eds. Digital soil mapping with limited data. Elsevier,Amsterdam,2008:15 - 30.

[12] 盛建东. 干旱区盐渍化土壤数字制图方法研究与应用[D]. 北京:中国农业大学,2009. [SHENG Jiandong. Research and application of digital mapping method for saline soil in arid area [D]. Beijing:China Agricultural University,2009.]

[13] 芦园园,张甘霖,赵玉国,等. 复杂景观环境下土壤厚度分布规则提取与制图[J]. 农业工程学报,2014,30(18):132 - 141. [LU Yuanyuan,ZHANG Ganlin,ZHAO Yuguo,et al. Extracting and mapping of soil depth distribution rules in complex landscape environment[J]. Transactions of the Chinese Society of Agricultural Engineering,2014,30(18):132 - 141.]

[14] 朱阿兴,李宝林,杨琳,等. 基于 GIS、模糊逻辑和专家知识的土壤制图及其在中国应用前景[J]. 土壤学报,2005,42(5):844 - 851. [ZHU Axing,LI Baolin,YANG Lin,et al. Predictive soil mapping based on a GIS,expert knowledge,and fuzzy logic framework and its application prospects in China[J]. Acta Pedologica Sinica,2005,42(5):844 - 851.]

[15] 朱阿兴,李宝林,裴韬,等. 精细数字土壤普查模型与方法[M]. 北京:科学出版社,2008:117 - 128. [ZHU Axing,LI Baolin,PEI Tao,et al. Model and method of fine digital soil census [M]. Beijing:Science Press,2008:117 - 128.]

[16] ZHU A X,YANG L,LI B,et al. Construction of membership functions for predictive soil mapping under fuzzy logic[J]. Geoderma,2010,155(3 - 4):166 - 174.

[17] 杨琳,朱阿兴,秦承志,等. 运用模糊隶属度进行土壤属性制图的研究——以黑龙江鹤山农场研究区为例[J]. 土壤学报,2009,46(1):9 - 14. [YANG Lin,ZHU Axing,QIN Chengzhi,et al. Soil property mapping using fuzzy membership:A case study of a study area in Heshan Farm of Heilongjiang Province[J]. Acta Pedologica Sinica,2009,46(1):9 - 14.]

[18] 杨琳,朱阿兴,秦承志,等. 基于典型点的目的性采样设计方法及其在土壤制图中的应用. 地理科学进展,2010,29(3):279 - 286. [YANG Lin,ZHU Axing,QIN Chengzhi,et al. A purposive sampling design method based on typical points and its application in soil mapping[J]. Progress in Geography,2010,29(3):279 - 286.]

[19] 王改粉,赵玉国,杨金玲,等. 流域尺度土壤厚度的模糊聚类与预测制图研究[J]. 土壤,2011,43(5):835 - 841. [WANG Gaifen, ZHAO Yuguo, YANG Jinling,et al. Prediction and mapping of soil depth at a watershed scale with fuzzy-c-means clustering method[J]. Soils,2011,43(5):835 - 841.]

[20] 卢岩君,秦承志,邱维理,等. 基于少量典型样点土壤属性空间分布推测模型中的土壤属性参数敏感性分析[J]. 地理科学,2011,31(12):1549 - 1554. [LU Yanjun,QIN Chengzhi,QIU Weili,et al. Sensitivity analysis of soil property parameter in typical-sample-based prediction model of digital soil mapping[J]. Scientia Geographica Sinica,2011,31(12):1549 - 1554.]

[21] 杨琳,朱阿兴,秦承志,等. 一种基于样点代表性等级的土壤采样设计方法[J]. 土壤学报,2011,48(5):938 - 946. [YANG Lin,ZHU Axing,QIN Chengzhi,et al. A soil sampling method based on representativeness grade of sampling points[J]. Acta Pedologica Sinica,2011,48(5):938 - 946.]

[22] 刘京,朱阿兴,张淑杰,等. 基于样点个体代表性的大尺度土壤属性制图方法[J]. 土壤学报,2013,50(1):12 - 20. [LIU Jing,ZHU Axing,ZHANG Shujie,et al. Large scaled soil attribute mapping method based on individual representativeness of sample sites [J]. Acta Pedologica Sinica,2013,50(1):12 - 20.]

[23] 毛丽丽,于静洁,张一驰. 模糊 c 均值聚类方法在黑河下游土壤属性制图中的初步应用研究[J]. 干旱区资源与环境,2013,27(1):195 - 201. [MAO Lili,YU Jingjie,ZHANG Yichi. Preliminary application of the fuzzy c-mean clustering method to the soil mapping in the lower reaches of Heihe River[J]. Journal of Arid Resources and Environment,2013,27(1):195 - 201.]

[24] 李润奎,彭明,河野泰之,等. 基于分层策略和模糊推理的北方石质山区土壤厚度制图——以滦平县虎什哈流域为例[J]. 地理研究,2013,32(5):965 - 973. [LI Runkui,PENG Ming,YASUYUKI Kono,et al. Soil depth mapping in lithoidal mountainous area using stratified strategy and fuzzy logic:A case study in Hushiha watershed,North China[J]. Geographical Research,2013,32(5):965 - 973.]

[25] 张淑杰,朱阿兴,刘京,等. 基于样点的数字土壤属性制图方法及样点设计综述[J]. 土壤,2012,44(6):917 - 923. [ZHANG Shujie,ZHU Axing,LIU Jing,et al. Sample-based digital soil mapping methods and related sampling schemes[J]. Soils,2012,44(6):917 - 923.]

[26] YANG L,ZHU A X,QI F,et al. An integrative hierarchical stepwise sampling strategy and its application in digital soil mapping [J]. International Journal of Geographical Information Science,2013,27(1):1 - 23.

[27] 肖笃志,胡玉昆. 天山南坡巴音布鲁克盆地主要退化草场调查[J]. 中国草地,1991,(4):40 - 44. [XIAO Duzhi,HU Yukun. Investigation of the main degenerate grassland in the Bayanbulak Basin in Tianshan Mountain [J]. Grassland of China,1991,(4):40 - 44.]

[28] 巴音郭楞蒙古自治州概况编委会. 巴音郭楞蒙古自治州概况[M]. 北京:民族出版社,2009:197 - 198. [The Editorial Committee of General Situation of Bayingol Mongol Autonomous Prefecture. General situation of Bayingol Mongol Autonomous Prefecture [M]. Beijing:The Ethnic Publishing House,2009:197 - 198.]

[29] 杜军剑,张仕明,李刚. 巴音布鲁克草原植被生长气象条件指数变化特征[J]. 沙漠与绿洲气象,2014,8(1):45 - 50. [DU Junjian,ZHANG Shiming,LI Gang. The meteorological conditions index variation characteristics of grassland vegetation growth over

Bayanbulak[J]. Desert and Oasis Meteorology,2014,8(1):45－50.]

[30] 王鑫,胡玉昆,热合木都拉·阿迪拉,等. 巴音布鲁克高寒草原羊茅(*Festuca ovina*)群落生态位特征[J]. 干旱区地理,2009,32(2):255－260.[WANG Xin,HU Yukun,ADILA Rehemudula,et al. Niche characteristics of *Stipa purpurea* community in alpine meadow of Bayanbulak[J]. Arid Land Geography,2009,32(2):255－260.]

[31] THOMPSON J A,BELL J C,BUTLER C A. Digital elevation model resolution:Effects on terrain attribute calculation and quantitative soil-landscape modeling[J]. Geoderma,2001,100(1－2):67－89.

[32] WANG J,FU B J,QIU Y,et al. Soil nutrients in relation to land use and landscape position in the semi-arid small catchment on the Loess Plateau in China[J]. Journal of Arid Environments,2001,48(4):537－550.

[33] SEIBERT J,STENDAHL J,SΦRENSEN R. Topographical influences on soil properties in boreal forests[J]. Geoderma,2007,141(1－2):139－148.

[34] MCSWEENEY K,SLATER B K,HAMMER R D,et al. Towards a new framework for modeling the soil-landscape continuum[M]//AMUNDSON R,ed. Factors of soil formation:A fiftieth anniversary publication. Madison,Wisconsin:Soil Science Society of America,1994:127－154.

[35] 黄昌勇,徐建明. 土壤学[M]. 北京:中国农业出版社,2010:87－88.[HUANG Changyong,XU Jianming. Pedology[M]. Beijing:China Agriculture Press,2010:87－88.]

[36] SUN X L,ZHAO Y G,WANG H L,et al. Sensitivity of digital soil maps based on FCM to the fuzzy exponent and the number of clusters. Geoderma,2012,171－172:24－34.

[37] BEZDEK J C,EHRLICH R,FULL W. FCM:The fuzzy c-means clustering algorithm[J]. Computers and Geosciences,1984,10(2/3):191－203.

[38] 杜薇,王昌全,李冰,等. 基于GIS和地统计学的攀枝花烟区土壤pH空间变异特征研究[J]. 核农学报,2012,26(6):924－929.[DU Wei,WANG Changquan,LI Bing,et al. Spatial variability of soil pH based on GIS combined with geostatistics in Panzhihua tobacco area[J]. Journal of Nuclear Agricultural Sciences,2012,26(6):924－929.]

[39] 谭艳,吴承祯,洪伟,等. 邓恩桉林地土壤pH空间变异分析[J]. 植物资源与环境学报,2012,21(1):14－19.[TAN Yan,WU Chengzhen,HONG Wei. Analysis on spatial variation of pH in soil of *Eucalyptus dunnii* forest[J]. Journal of Plant Resources and Environment,2012,21(1):14－19.]

[40] ZHU A X,BAND L E,VERTESSY R,et al. Derivation of soil properties using a soil land inference model(SoLIM)[J]. Soil Science Society of America Journal,1997,61:523－533.

[41] 王库. 地理权重回归在土壤pH空间预测中的应用[J]. 湖南农业大学学报(自然科学版),2013,39(1):73－79.[WANG Ku. Application of geographically weighted regression on the spatial prediction of soil pH[J]. Journal of Hunan Agricultural University(Natural Science Edition),2013,39(1):73－79.]

[42] 刘琼峰,李明德,段建南,等. 农田土壤铅、镉含量影响因素地理加权回归模型分析[J]. 农业工程学报,2013,29(3):225－234.[LIU Qiongfeng,LI Mingde,DUAN Jiannan,et al. Analysis on influence factors of soil Pb and Cd in agricultural soil of Changsha suburb based on geographically weighted regression model[J]. Transactions of the Chinese Society of Agricultural Engineering,2013,29(3):225－234.]

[43] 瞿明凯,李卫东,张传荣,等. 地理加权回归及其在土壤和环境科学上的应用前景[J]. 土壤,2014,46(1):15－22.[QU Mingkai,LI Weidong,ZHANG Chuanrong,et al. Geographically weighted regression and its application prospect in soil and environmental sciences[J]. Soils,2014,46(1):15－22.]

[44] 杨顺华,张海涛,陈家赢,等. 平原丘陵过渡带土壤有机碳空间分布及环境影响[J]. 中国环境科学,2015,35(12):3728－3736.[YANG Shunhua,ZHANG Haitao,CHEN Jiaying,et al. The spatial variability of soil organic carbon in plain-hills transition belt and its environmental impact[J]. China Environmental Science,2015,35(12):3728－3736.]

[45] FISCHER M M,GETIS A. Handbook of applied spatial analysis[M]. Heidelberg,Dordrecht,London and New York:Springer,2010:461－486.

[46] 周启鸣,刘学军. 数字地形分析[M]. 北京:科学出版社,2006:13－25.[ZHOU Qiming,LIU Xuejun. Digital terrain analysis[M]. Beijing:Science Press,2006:13－25.]

[47] 汤国安,刘学军,房亮,等. DEM及数字地形分析中尺度问题研究综述[J]. 武汉大学学报:信息科学版,2006,31(12):1059－1066.[TANG Guoan,LIU Xuejun,FANG Liang,et al. A review on the scale issue in DEMs and digital terrain analysis[J]. Geomatics and Information Science of Wuhan University,2006,31(12):1059－1066.]

[48] 刘学军,卢华兴,仁政,等. 论DEM地形分析中的尺度问题[J]. 地理研究,2007,26(3):433－442.[LIU Xuejun,LU Huaxing,REN Zheng,et al. Scale issues in digital terrain analysis and terrain modeling[J]. Geographical Research,2007,26(3):433－442.]

[49] 呼雪梅,秦承志. 地形信息对确定DEM适宜分辨率的影响[J]. 地理科学进展,2014,33(1):50－56.[HU Xuemei,QIN Chengzhi. Effects of different topographic attributes on determining appropriate DEM resolution[J]. Progress in Geography,2014,33(1):50－56.]

Predictive soil pH mapping based on fuzzy clustering in typical alpine grassland of Xinjiang

ZHU Lei[1,2], SHENG Jian-dong[1,2], JIA Hong-tao[1,2]

(1 *College of Grassland and Environment Sciences, Xinjiang Agricultural University, Urumqi* 830052, *Xinjiang, China*;
2 *Xinjiang Key Laboratory of Soil and Plant Ecological Processes, Urumqi* 830052, *Xinjiang, China*)

Abstract: Obtaining spatial distribution of grassland soil properties accurately and efficiently can provide basic information and reference for grassland resource management. Compared with the traditional soil survey method, the soil environment inference model based on fuzzy logic can improve the efficiency of field sampling and accuracy of predictive mapping, which has been widely used in digital soil mapping. However, due to the soil's spatial variation and its non-linearity with the environmental conditions, the stability of the existing models are relatively low. The models have not been applied in alpine meadow area. In this study, fuzzy C-means clustering (FCM) was used to predict soil pH in the surface layer of grassland soil within a 4 km^2 area in Bayanbulak District, Xinjiang Uyghur Autonomous Region, China. Six terrain factors, including elevation, slope, aspect, planform curvature, profile curvature and topographic wetness index, were clustered. Fuzzy membership of 9 groups of environmental factors were derived to position 18 soil samples in the area with membership larger than 0.9. Then pH distribution was predicted with fuzzy membership model. The pH value of study ranged from 7.170 to 8.186 and was consistent with the measured values. The mapping results reflected continuous changing of soil properties with terrain changing. There were 35 individual soil samples (16 equal-interval sampling points, 9 cross-sectional sampling points and 10 sampling points according to altitude) collected as validation data set. The agreement coefficients between observed values and predicted values were high, and the accuracy of FCM model is higher than that of Ordinary Kriging method and Inverse Distance Weighted method. FCM and purposive sampling for digital soil mapping is also suitable for small-scale region. This approach is an efficient digital soil mapping method with satisfactory prediction precision using less samples. It could be possibly applied to the areas with the similar landscape conditions.

Key words: fuzzy clustering; purposive sampling; pH; Bayanbulak